

PATENT  
Atty. Dkt. No. 2002-0447

## **IN THE CLAIMS**

This listing of claims will replace all prior versions, and listings, of claims in the application:

Claims 1-3 (Canceled).

4. (Currently Amended) ~~The method of Claim 3~~ A method of searching a document having nested-structure document-specific markup, the method comprising:

receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process;

deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the step of deriving the query-specific indices includes forming at least one of a group including:

an index of each word in the phrase to be matched by the phrase matching process;

an index of context tags that may be found in the document; and

an index of at least a tag or annotation to be ignored during the phrase matching process,

wherein the query-independent indices were created by a method including:

a) labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the

PATENT  
Atty. Dkt. No. 2002-0447

document, a position in the document, and by then indicating that the word or tag is present or not present at that position,

wherein the step of deriving the query-specific indices involves deriving the query-specific indices from the query-independent indices without rebuilding any of the query-independent indices; and

carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

5. (Currently Amended) ~~The method of Claim 2~~ A method of searching a document having nested-structure document-specific markup, the method comprising:

receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process, wherein the phrase matching process includes:

for each context interval, defined by a beginning index defining a position of beginning tag and a closing index defining a position of a closing tag, performing an index-nested loop by probing an index of each phrase word in order, and an index of each tag or annotation to be ignored, so as to construct at least one witness;

wherein each witness is a contiguous sequence of intervals contained within the context interval and includes each phrase word occurrence exactly once and in phrase order;

deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the query-independent indices were created by a method including:

a) labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

PATENT  
Atty. Dkt. No. 2002-0447

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating that the word or tag is present or not present at that position; and

carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

6. (Original) The method of Claim 5, wherein at least one witness includes each phrase word occurrence exactly once and in phrase order, interleaved with tags or annotations to be ignored.

7. (Currently Amended) ~~The method of Claim 2~~ A method of searching a document having nested-structure document-specific markup, the method comprising:

receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process, wherein the phrase matching process includes:

scanning, in document order, a combined index of (A) phrase words and (B) tags or annotations to be ignored, while using a stack to keep track of nested context intervals and annotation intervals;

wherein:

the stack includes at least one entry corresponding to a current context interval in which witnesses are identified; and

the at least one entry maintains a set of (A) partial witnesses that are being identified and (B) complete witnesses that have been identified, within the current context interval;

PATENT  
Atty. Dkt. No. 2002-0447

deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the query-independent indices were created by a method including:

a) labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating that the word or tag is present or not present at that position; and

carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

Claims 8-13 (Cancelled)

14. (Currently Amended) ~~The computer program product of claim 13~~ A computer program product including computer executable code or computer executable instructions that, when executed, causes a computer to perform a method of searching a document having nested-structure document-specific markup, the method comprising:

receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process;

deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the query-independent indices were created by a method including:

a) labeling elements in the document with intervals, wherein:

PATENT  
Atty. Dkt. No. 2002-0447

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating whether or not the word or tag is present at that position,

wherein the phrase matching process includes:

for each context interval, defined by a beginning index defining a position of beginning tag and a closing index defining a position of a closing tag, performing an index-nested loop by probing an index of each phrase word in order, and an index of each tag or annotation to be ignored, so as to construct at least one witness;

wherein each witness is a contiguous sequence of intervals contained within the context interval and includes each phrase word occurrence exactly once and in phrase order; and

carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

15. (Currently Amended) ~~The program product of claim 13~~ A computer program product including computer executable code or computer executable instructions that, when executed, causes a computer to perform a method of searching a document having nested-structure document-specific markup, the method comprising:

receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process;

PATENT  
Atty. Dkt. No. 2002-0447

deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the query-independent indices were created by a method including:

a) labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating whether or not the word or tag is present at that position, wherein the phrase matching process includes:

scanning, in document order, a combined index of (A) phrase words and (B) tags or annotations to be ignored, while using a stack to keep track of nested context intervals and annotation intervals;

wherein:

the stack includes at least one entry corresponding to a current context interval in which witnesses are identified; and

the at least one entry maintains a set of (A) partial witnesses that are being identified and (B) complete witnesses that have been identified, within the current context interval; and

carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

16. (Canceled)

17. (Canceled)

PATENT  
Atty. Dkt. No. 2002-0447

18. (Currently Amended) ~~The system of claim 17~~ A system for searching a document having nested-structure document-specific markup, the system comprising:

means for receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process;

means for deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the means for deriving query-independent indices comprises:

a) means for labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) means for forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating whether or not the word or tag is present at that position,

wherein the phrase matching process includes:

for each context interval, defined by a beginning index defining a position of beginning tag and a closing index defining a position of a closing tag, means for performing an index-nested loop by probing an index of each phrase word in order, and an index of each tag or annotation to be ignored, so as to construct at least one witness;

wherein each witness is a contiguous sequence of intervals contained within the context interval and includes each phrase word occurrence exactly once and in phrase order; and

PATENT  
Atty. Dkt. No. 2002-0447

means for carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

19. (Currently Amended) ~~The system of claim 17~~ A system for searching a document having nested-structure document-specific markup, the system comprising:

means for receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process;

means for deriving query-specific indices based on query-independent indices that were created specific to each document, wherein the means for deriving query-independent indices comprises:

a) means for labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) means for forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating whether or not the word or tag is present at that position, wherein the phrase matching process includes:

means for scanning, in document order, a combined index of (A) phrase words and (B) tags or annotations to be ignored, while using a stack to keep track of nested context intervals and annotation intervals;

wherein:



PATENT  
Atty. Dkt. No. 2002-0447

the stack includes at least one entry corresponding to a current context interval in which witnesses are identified; and

the at least one entry maintains a set of (A) partial witnesses that are being identified and (B) complete witnesses that have been identified, within the current context interval; and

means for carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

20. (Canceled).